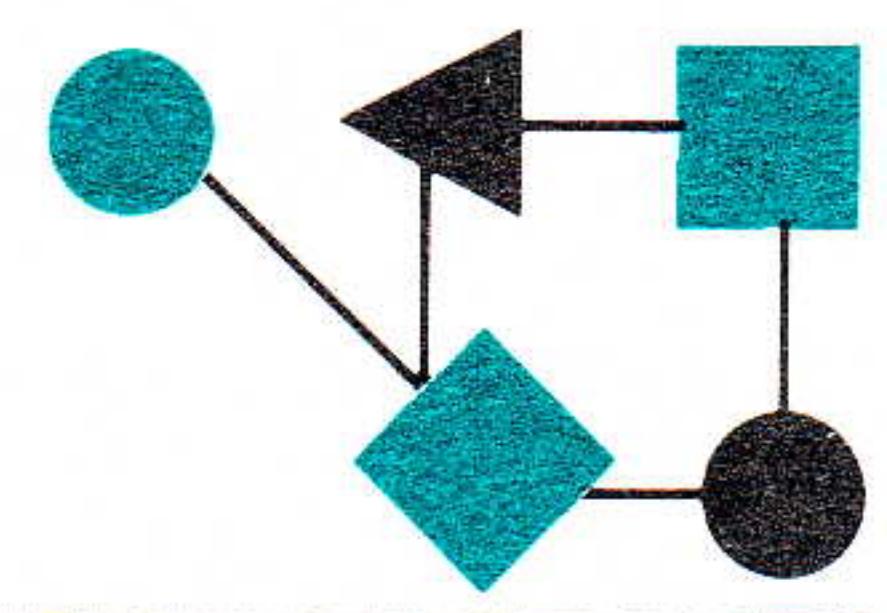


CONNE~~X~~IONS



The Interoperability Report

November 1988

Volume 2, No. 11

ConneXions — The Interoperability Report tracks current and emerging standards and technologies within the computer and communications industry.

In this issue:

Networking in Canada.....	2
OSI Conference.....	6
INTEROP 88 report.....	7
The IETF Host Requirements Working Group.....	10
Improving Your TCP: The Checksum Routine.....	14
Recent Publications.....	15

ConneXions is published by Advanced Computing Environments, Inc., 480 San Antonio Road, Suite 100, Mountain View, CA 94040, USA. Phone: 415-941-3399

© 1988
Advanced Computing Environments.
Quotation with attribution encouraged.

ISSN 0894-5926

From the Editor

This month we have a brief report from INTEROP 88, our largest conference to date. In future issues of *ConneXions*, there will be articles on topics covered in the conference, in particular we hope to bring you an inside look at what it was like to bring up a show internet in a few hectic days. In response to many request, a simplified map of the Show and Tel-Net appears on page 8.

In our series looking at existing or evolving computer networks, we come to Canada. Philip Prindeville of McGill University gives an overview of the Canadian networks on page 2.

Our next main event is the *OSI Product Integration Conference* which we are hosting jointly with COS at the end of this month. By now you should have received the Advance Program for this conference. A brief description appears on page 6.

Last year, RFC 1009 entitled "Requirements for Internet Gateways" was published. Soon there will be a companion document outlining the requirements for Internet Hosts. Craig Partridge and Bob Braden give an overview of this important document.

In our series of hints for TCP implementors and maintainers, Craig Partridge looks at the "tricks" you can do when computing the Internet Checksum.

Since our July issue where we listed sources of information on TCP/IP, a number of new and relevant publications have been issued. Page 15 has an overview of these documents.

I am still looking for input from the *ConneXions* readership, suggestions for topics, user experiences, horror stories or questions for the gurus, etc. Please give me a call or drop a line. Your input is very much appreciated.

All attendees of the 2nd TCP/IP Interoperability Conference which was held last December received a free one-year subscription to *ConneXions*. These subscriptions will run out soon, and we encourage you to renew to ensure continued uninterrupted service. A renewal reminder should arrive within the next month. Make sure you take the appropriate renewal steps. 1989 promises to be an exciting year with many interesting articles in the field of interoperability, stay tuned!

Pleins feux sur les réseaux canadiens (Networking in Canada)

by Philip A. Prindeville, McGill University

Overview

Canada has been networking for several years. Until the present, it has been an *ad hoc* collection of private groups supporting individual interests. Today, however, there is a growing need to communicate between these groups and share resources, and to actively move (through usage and participation in the standards process) towards true Open Systems Interconnection.

NetNorth

NetNorth is a national private network connecting academic and research sites all over Canada (see map on page 5), and is similar in design and scope to BITNET or EARN. The network is composed of 56 member organizations (51 are educational, 1 commercial, and 4 government) with 166 nodes, linked via leased lines between 2.4 and 9.6kbps. NetNorth has connections to BITNET, EARN, and ASIAnet. The network uses the IBM Network Job Entry protocol (NJE) to provide file transfer services, on top of which remote execution and electronic mail is layered. Many non-IBM systems are on the network, including Amdahl, CDC, DEC, Honeywell, Prime, and Sun.

Commercial Services

Several networking services are offered by Telecom Canada, a consortium of the regional telephone operating companies: Envoy 100, an electronic message service, and Datapac, which was the first commercial X.25 public data network (PDN) in the world. Most of these carriers also offer 56kbps and T1 (1.5Mbps) facilities, and ISDN is beginning trial service.

CDNnet

An alternative electronic messaging (and file transfer) service is offered through the CDNnet, headquartered at the University of British Columbia (UBC). It is the world's first X.400 network. Membership is available to any Canadian organization involved in education, research, or advanced development; currently there are 23 educational, 6 government, 2 commercial, and 2 non-profit sites. It provides delivery (or non-delivery) notification, directory service, and gateways to non-X.400 networks such as the Internet, CSNET, BITNET, and UUCP, and software is included with membership. Interconnection to other member sites is via X.25 PDN circuits. The UBC gateway currently passes approximately 5,500 messages a day to/from other networks.

The Internet

A few Canadian sites which have been cooperating on research projects with U.S. counterparts are connected to the Arpanet, including sites belonging to the Directorate Radio Propagation and Systems Communications Research Centre, and the Defense Research Establishment Atlantic. Within the last year, a substantial amount of interest in internetworking has developed at universities. As a result, a few regional networks have connected to the Internet.

ONet

The Ontario Network is a consortium of several universities—Toronto, Western Ontario, Queen's, York, McMaster, and Waterloo, with Guelph and Carleton possibly joining later—and government research facilities—Information Technology Research Centre and Institute for Space and Terrestrial Science—connected via TCP/IP and DECnet by cisco routers over 19.2kbps leased lines, with 56kbps discussed for the future.

(By the time this goes to press, all links should be operational.) The network is supported by membership dues and contributions from the Ontario Centre for Large Scale Computing. Connection to the Internet is via a 56kbps line to Cornell, and there is discussion about participating as a BITNET II pilot site sharing this same facility.

Québec

CRIMnet is a Montréal area network designed to serve its member institutions (Centre de Recherche Informatique de Montréal (CRIM), McGill University, Concordia University, Université de Montréal, Université du Québec à Montréal, and École Polytechnique). The network is based on DECnet routers and 56kbps leased lines, so unfortunately most traffic—TCP/IP—must be encapsulated (using a local scheme). Right now there is a great deal of discussion regarding extending the network to other Québec universities (such as Laval, Sherbrooke, and the various campuses of Université du Québec), upgrading the lines to higher speeds and running TCP/IP natively, and providing access to new supercomputing facilities being planned in Montréal. McGill maintains a 9.6kbps leased line TCP/IP connection to BBN in Cambridge and CRIM has plans to connect as well, sometime in November.

BCnet

The universities of British Columbia at Vancouver (UBC), Victoria, and Simon Fraser, with the TRIUMF Cyclotron, Dominion Astrophysical Observatory (NRC) government laboratories and Microtel Pacific Research Corporation have built a wide area network using bridged Ethernet, Vitalink repeaters, and Proteon routers, and links from 56kbps to T1. Many protocols are supported including TCP/IP, X.25, and DECnet. UBC has been providing an Internet connection since late summer via NorthWestnet in Seattle. All other regions in Canada are planning to implement similar networks.

National Research Council (NRC)

Canada has an extensive scientific community that is supported by the federal government. As an example, the NRC's Canada Institute for Scientific and Technical Information provides databases in Agriculture, Biology, Chemistry, Education, Geology, Medicine, Physics and Engineering, Metallurgy, and Social Sciences. These are available to commercial and university users, and are accessible in several ways including post, commercial E-mail services (such as Envoy), dial-up lines, and dedicated terminals. Some of these databases are the most extensive of their type in the world.

The Prime Minister recently announced the "Centres of Excellence" programme to fund distributed research groups of outstanding calibre. This is intended to encourage cooperation amongst industry and academia in areas that are of strategic interest to Canada.

NRCnet

Last year the NRC proposed a national research network (NRCnet), and with the announcement of the Centres of Excellence it became evident that the programme would be well-served by this network. It is to be a *production* network in support of research rather than an experimental one. Nonetheless it is clear that networking research will also be one of the areas affected by this project, and perhaps a subset of the network can be allocated for development.

Networking in Canada (*continued*)

The main purpose of the network is to provide a means for researchers to communicate with each other easily and efficiently, and to access specialized facilities such as supercomputing centres. The network will be like the National Science Foundation's network (NSFnet) in that it will consist of a coast-to-coast communications backbone, with regional networks being attached at strategic hubs. The map (page 5) shows an example topology.

Proposals

Already, two working papers for the switching fabric have been submitted. The first was suggested by the University of Toronto and involved IBM Canada. This was quite similar to the NSFnet with the backbone being a loosely coupled network of IBM PC/RT operating as switches, and dedicated trunks of 56kbps or T1. Later development may include multi-protocol support. The chief strengths of this design are the use of existing technology, and a base level of functionality for all users.

The other paper was from UBC, and called for a high-speed (T1) X.25 subnet of an undisclosed design. The advantage of this approach is that it provides a communications subnet to the various existing parties involved in networking (TCP/IP, X.400, and DECnet can all be run atop X.25). The NRC will soon issue a Request for Proposals (RFP) for network operation, which it will support through contributions. It is hoped that a pilot implementation will be ready before the year's end. Funding will come in part from the federal government, with self-sufficiency expected in 5 years.

Issues

There are many issues involved in designing a network that is optimal for Canada. One of the most apparent problems is Canada's size: though it has a larger area than the U.S., its population is much less, and this increases long distance costs (indeed, one investigation revealed it was cheaper to run lines south to the U.S. and then cross-continent). Nor is there the mixed blessing of divestiture. The regulatory environment is much more restrictive than in the U.S.

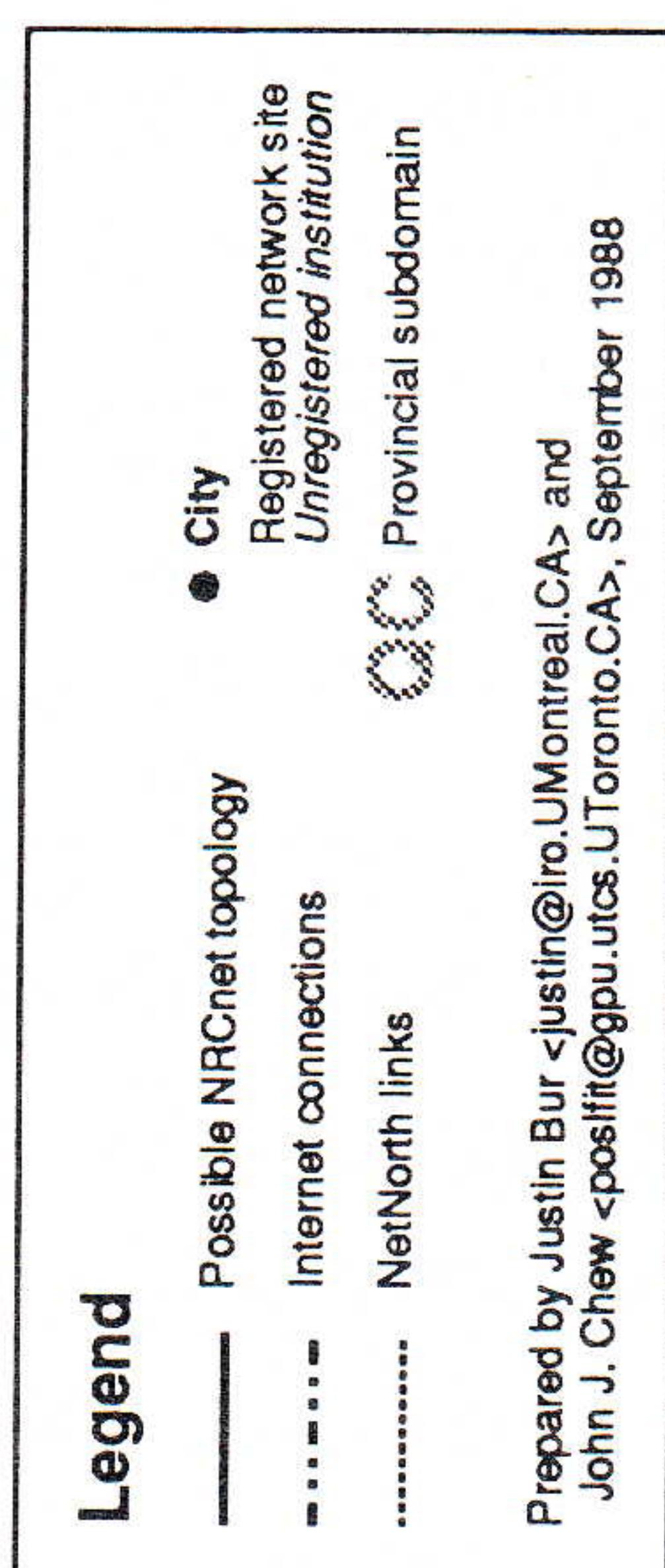
As indicated above, different camps would be better served by one architecture over another. Physicists tend to prefer DECnet (SPAN—Space/Physics Analysis Network, NASA, and HEPnet—High Energy Physics Network, U.S. DoE). CDNnet is well served by something that X.400 can be run on top of, such as X.25. NetNorth, through recent work on BITNET, can run on top of TCP/IP. It is the general consensus that TCP/IP would serve all camps best, and provide a basic level of functionality. At the earliest feasible point in the future, NRCnet will begin transition to an ISO suite.

Lastly, there is policy-based routing, which is still being developed; clearly Canadian traffic should traverse Canada unless a backbone link fails. On the bright side, however, we have a large telecommunications industry and sufficient digital facilities linking metropolitan areas.

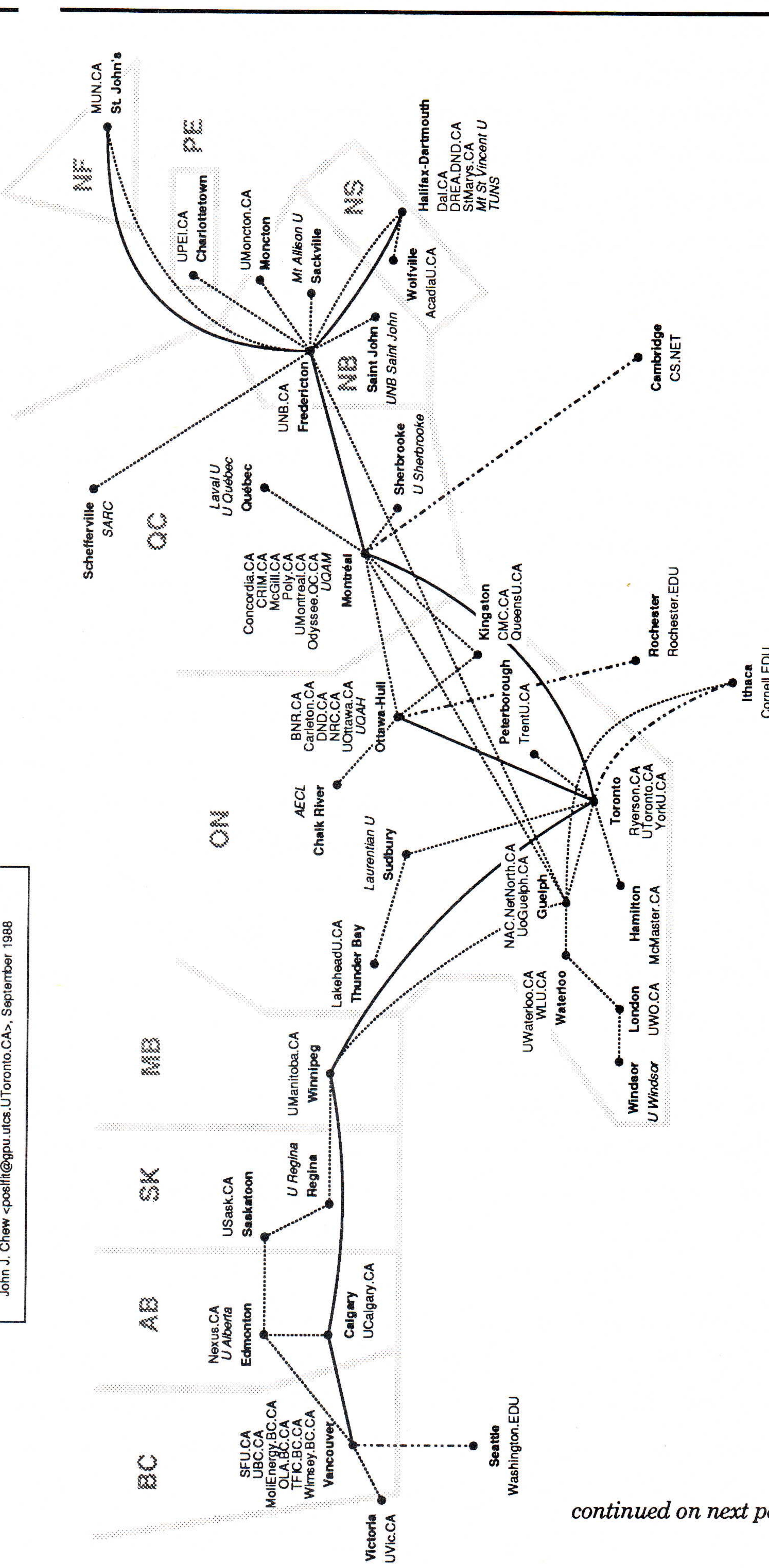
The CA Domain

Closely related to Internet connection and vital to E-mail exchange is domain registration. After some deliberation, it was decided to register Canada according to its ISO 3166 two-letter country code, which is CA.

Major Network Sites in Canada



Prepared by Justin Bur <justin@iro.UMontreal.CA> and
John J. Chew <positif@gpu.utcs.UToronto.CA>, September 1988



continued on next page

Networking in Canada (*continued*)

The registrar for the CA domain is John Demco of UBC: <CA-Registrar@RELAY.UBC.CA>. Currently, there are 40 domains registered: 30 educational, 10 commercial, and 2 government.

To register, request an application form by sending mail to <Archive-Server@RELAY.UBC.CA> with the subject (or message body) "send ca-domain application-form."

Further information

If you'd like more information about networking developments, send mail to <Listmaster@CS.McGill.CA> enumerating your interests.

Acknowledgements

The author wishes to thank the following persons for their assistance in completing this report: John Demco and Marilyn Martin of CDNnet, Roger Watt of the NetNorth Consortium, Nancy Fischer of the Network Information Center, Justin Bur of CRIM, John J. Chew at the University of Toronto, and Andy Woodsworth of the NRC Dominion Astrophysical Observatory. This report was based on information distributed via the NRC mailing list.

PHILIP A. PRINDEVILLE recently joined the staff of the McGill Research Centre for Intelligent Machines (MCRCIM) as a Research Associate, where he is involved in issues of design, implementation, and interoperability in heterogeneous networking environments. He has worked as a protocol implementor at FTP Software, Inc., and as a protocol developer at M.I.T./Project Athena and for Datapoint Corporation. Philip is a regular participant in Internet activities, chairs the IETF Synchronous Point-to-Point Standards Working Group, and contributes to the NRC network planning discussions.

OSI Product Integration Conference

Advanced Computing Environments and The Corporation for Open Systems (COS) will be sponsoring an *OSI Product Integration Conference* from November 29 through December 2nd, 1988. The conference will be held at the McLean Hilton outside Washington DC, 15 minutes from Dulles International Airport. The format is:

- *Two days of tutorials:* A total of 12 tutorials (some 1/2 day and some full day) are offered. Topics range from OSI fundamentals to Conformance Testing and Transition Strategies.
- *Technical sessions:* Over 30 technical sessions with particular emphasis on making OSI a reality through integration with existing systems. Issues covered include network management, The ISO-VTP Gateway, Government OSI profiles, X.400 for Users, Interoperability Testing, and OSI Security.
- *Vendor Implementation Descriptions:* Technical descriptions by vendors on their OSI offerings and development plans. Find out what OSI products are already available or on the horizon for your particular system.
- *Common Interest Sessions:* Similar to our normal BOFs, these sessions allow attendees to discuss OSI topics in an informal atmosphere. The sessions will be followed by a final plenary where a spokesperson from each common interest session will report its results.

For more information on this conference, call ACE at 415-941-3399.

INTEROP 88 Conference Report

Attendance

A total of 5529 people attended *INTEROP 88 The 3rd TCP/IP Interoperability Conference and Exhibition*, September 26—30, 1988 at the Santa Clara Convention Center and Doubletree Hotel. The conference and tutorials had an attendance of 1870, while 3659 were “exhibit-only” guests

Tutorials and BOFs

As in previous conferences, the tutorials on the first two days proved very popular, Doug Comer's *In-Depth Introduction to TCP/IP* attracted nearly 600 people for instance. In addition to the 16 technical session, the conference offered the opportunity for ad hoc meetings; Birds of a Feather sessions (BOFs). These sessions are clearly functioning as important venues for information exchange and we foresee the need to provide even more BOFs next time.

The Show and Tel-Net

For the first time in history, a “Show and Tel-Net” with over 50 vendors showing true TCP/IP interoperability was constructed in a little under 5 days. This major accomplishment was directed by Peter DeVries of the The Wollongong Group and Philip Almquist of Stanford University, with the help of a dozen or so “gurus” from the vendor/research community. I observed a wonderful cooperative spirit amongst these people who put in 20 hour workdays to make it all operational by the time the hall opened to the public.

The Shownet consisted of 9 different underlying network media, ranging from generic “yellow” Ethernet to fiber optics and T1 links. The T1 links provided us with connectivity to the Internet and allowed visitors to walk up to any terminal on the show floor, Telnet to their native system, and read their electronic mail. (See map, p8).

Network Management

As outlined in RFC 1052, the TCP/IP community has adopted the Simple Network Management Protocol (SNMP) as the short-term Internet standard. Additionally, a prototype network management system—“CMOT” or CMIP Over TCP/IP—has been designed by the NetMan group of the IETF. This system was being showcased at INTEROP 88 with 12 vendors participating. A number of SNMP implementations were also shown and new products announced. In future issues of *ConneXions*, we will explore both of these network management architectures in more detail.

On Thursday morning, September 29, the space shuttle Discovery lifted off, and I heard a few attendees wondering if they'll be able to contact the shuttle from next years show floor. I can just see it now: %ping discovery.shuttle.nasa.gov. Before the next INTEROP, the amateur packet radio community will have put the first IP host in space, perhaps we can ping that instead.

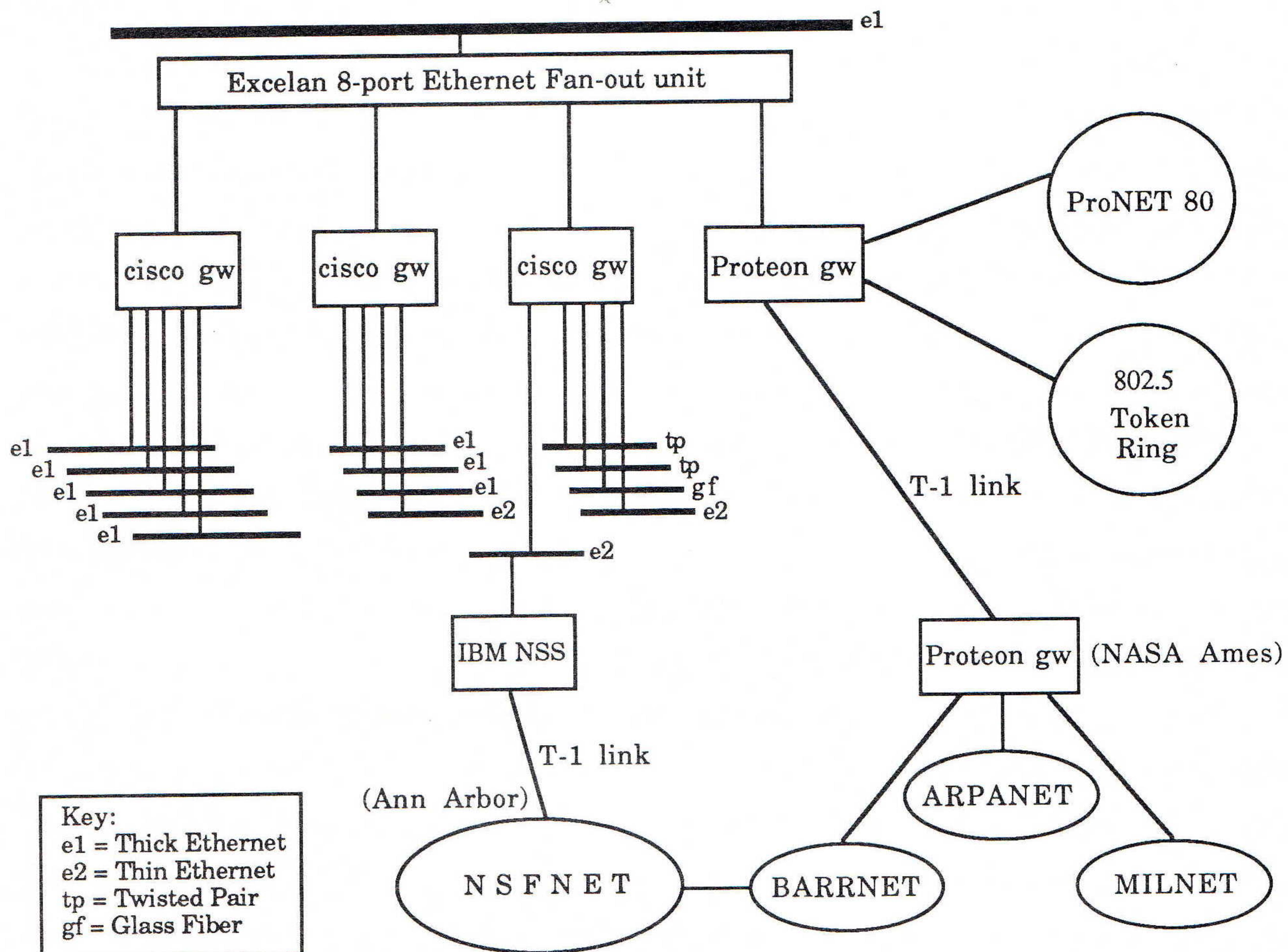
Cookies

We received a number of conference evaluation forms which we are in the midst of processing. Your comments are much appreciated and will help us in our planning for INTEROP 89. A number of people suggested that we publish the recipe for the Doubletree cookies, something we'd be happy to do were it not for the fact that this recipe is a well-kept trade secret. However, we promise to make every effort to have the cookies available at our future events.

Finally, our warmest thanks to all those who helped put together INTEROP 88, and thanks to all of our attendees, we'll see you next year in San Jose!

—Ole

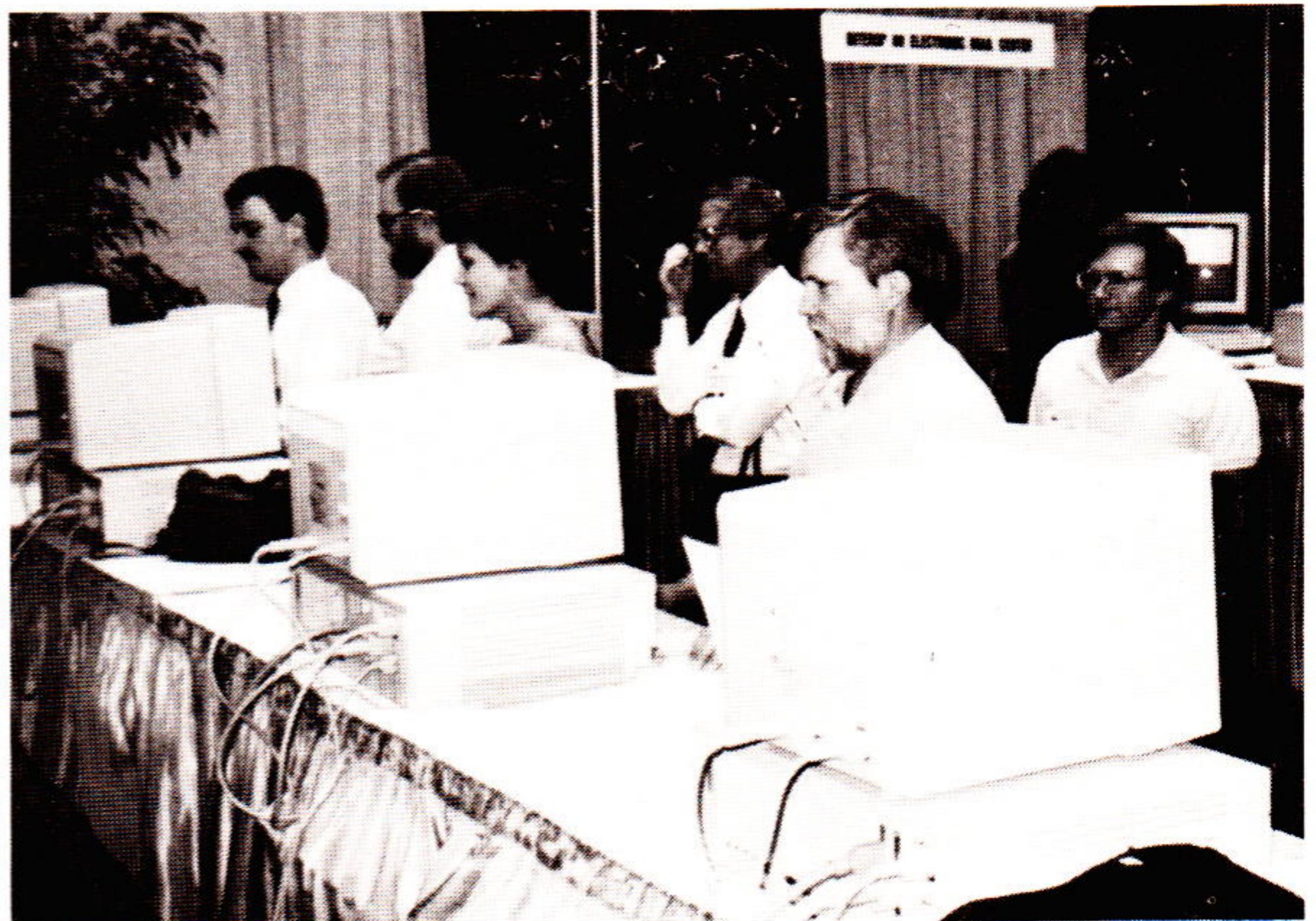
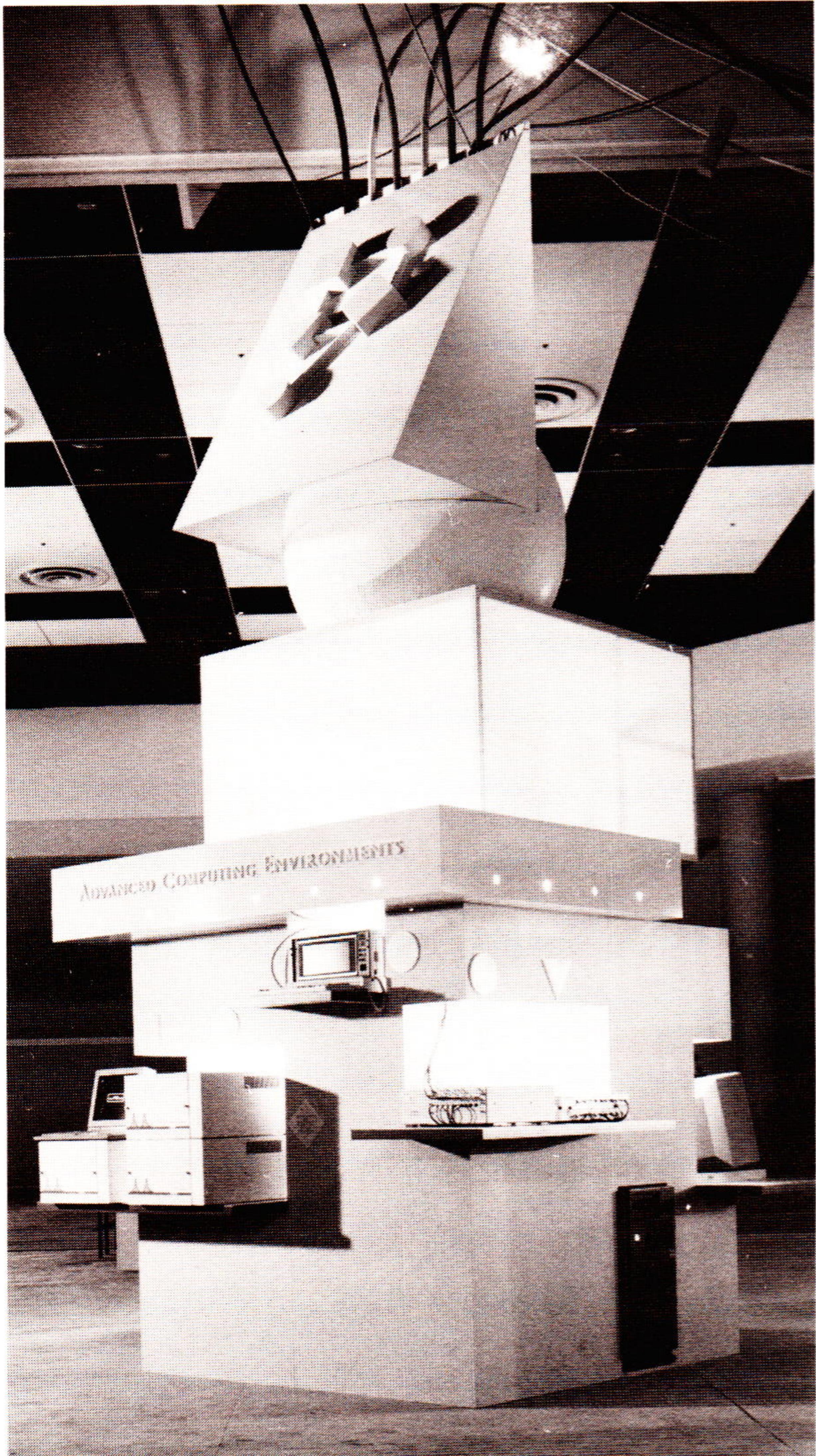
INTEROP 88

*Simplified Show and Tel-Net topology**Participating vendors:*

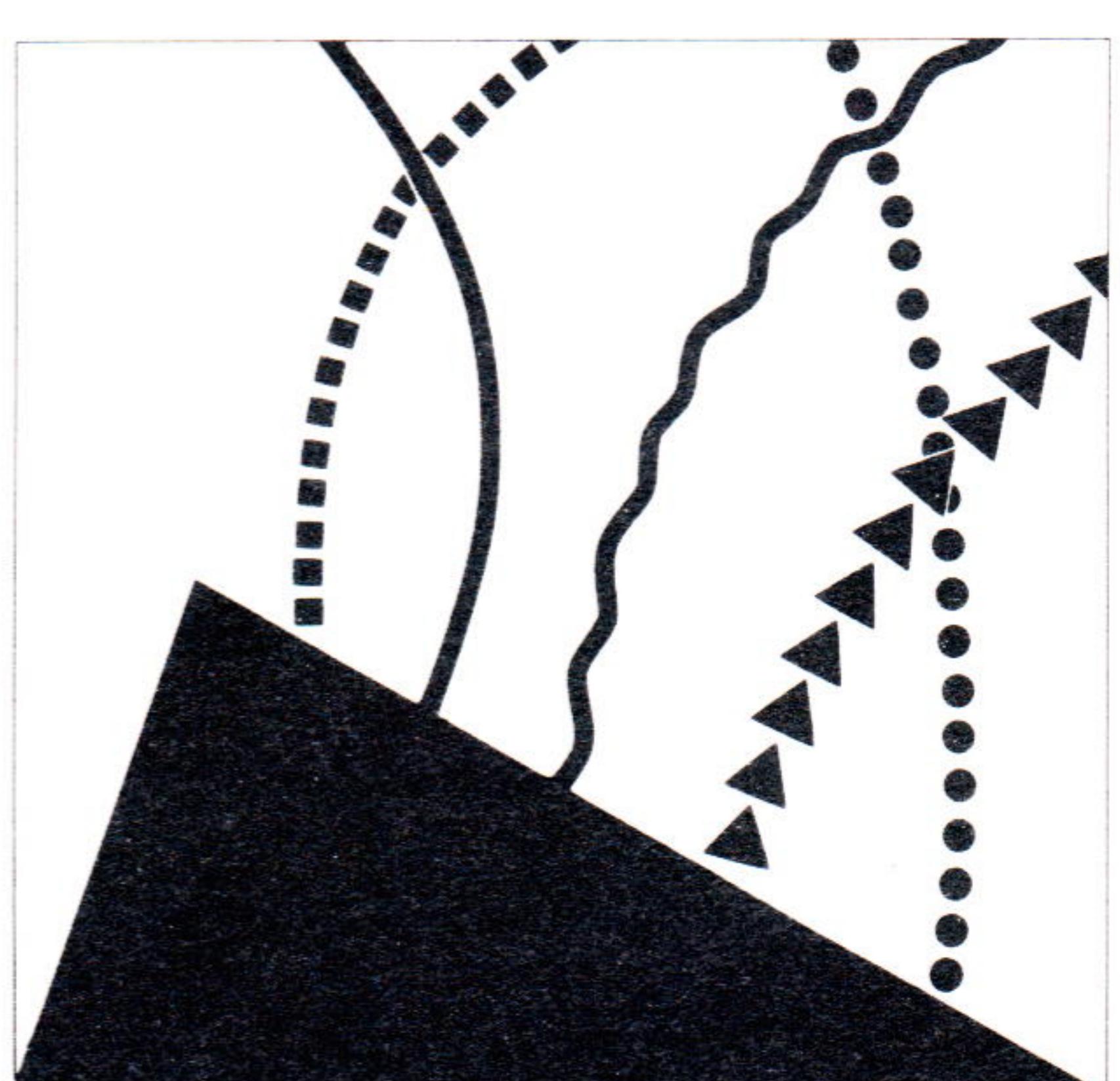
3Com
ACC
Apple Computer
Banyan Systems
BBN Communications
COMPUTERWORLD
CMC
Computer Network Technology
Concurrent Computer
Convergent Technologies
cisco Systems
DCA/SRI International
DEC
Encore
Eon Systems
Excelan/TGV/Kinetics
FTP Software

Halley Systems
Hewlett-Packard
Highland Software
IBM/MCI/Merit/CMU
Interactive Systems
InterCon
Interphase
Lachman Associates
Mitre/Unisys (NetMan)
Network General
Network Research
Network Solutions
Network Systems
Prentice-Hall
Prime Computer
Process Software
Proteon

Sirius Systems
Spider Systems
Sun Microsystems
SynOptics Communications
Syntax Systems/10Net
Sytek
Tandem Computers
TCL
TRW
Ungermann-Bass
UNIX World
Vitalink Communications
VXM Technologies/MIPS
Wellfleet Communications
Western Digital
The Wollongong Group
Xplex, Inc.



Above: The Network Operations Center. All cables (close to 3 miles in all) passed through here. Top right: Phil Karn, KA9Q demonstrates TCP/IP over an amateur packet radio channel. His station was linked to a similar setup in the Sirius Systems booth which allowed access to the Shownet "over the air." Above right: The Electronic Mail Center was a popular gathering place. Below: Doug Comer's TCP/IP tutorial attracted nearly 600 attendees.



INTEROP 88

3RD TCP/IP INTEROPERABILITY
CONFERENCE AND EXHIBITION

September 26 - 30, 1988

Santa Clara, California

Towards Full Interoperability— The IETF Host Requirements Working Group

by Craig Partridge and Bob Braden

Introduction

Last year the Internet community spent considerable time and effort putting together a specification for Internet gateways. This work resulted in the release of RFC 1009, "Requirements for Internet Gateways." [Ed.: See *ConneXions* Volume 1, No. 3, July 1987].

This year, the community has turned its attention to Internet *hosts*, and an Internet Engineering Task Force (IETF) working group, chaired by Bob Braden of ISI, has been hard at work writing a specification for Internet hosts. This working group, composed of experienced Internet researchers and implementors, has 30 members representing over a dozen organizations. Seven different authors have contributed protocol descriptions to the text.

Goals

The goals of the Host Requirements effort are much the same as those of RFC 1009. We would like to specify what we believe is required of a properly functioning IP host, and to clarify any ambiguities in the various RFCs that specify the protocols that hosts may need to support. In addition to the specific requirements, the document contains considerable supporting discussion, clarification, and commentary. We are trying to specify the best possible host implementation, not simply canonizing a particular implementation that the community feels is acceptable. In fact, it is the working group's opinion that no current implementation fully conforms the requirements document. At the same time, the Host Requirements document does *not* propose new Internet architecture. It is simply a much-needed specification of the implications for host software of the current architecture.

In general, the working group has striven to develop a requirement profile that allows a host to (1) interoperate correctly and efficiently with other conforming implementations; (2) interoperate with past implementations, including past mistakes; and (3) offers the hope of consistency with possible future Internet architectural changes.

To completely clarify all the protocols that could ever be used on an Internet host would be an enormous task. With over 1000 RFCs published there are simply too many protocols in use to discuss them all. Instead, the host requirements document looks at the core protocols and applications—those applications that any host should support.

This article gives an abbreviated overview of the current state of the host requirements document. The abbreviation must be severe; the current draft of the requirements document is now over 150 pages.

Organization

The document is organized by layers, from the link layer up to the application layer. Each major protocol has its own section. In general, each section begins with an introduction explaining the role of the protocol, followed by a "walk-through" of the RFC(s) defining the protocol. This "walk-through" corrects known errors and discusses ambiguities and controversial points in the protocol specification.

The walk-through is followed by a discussion of any issues not immediately raised by the specifications (for example, the description of *TCP Silly Window Syndrome avoidance* algorithms comes here, because this subject was not dealt with in RFC 793). The section concludes with a discussion of how the protocol must interface to the next higher layers, i.e., the “service interface,” and a list of the key documents that an implementor must read.

Terminology

The working group has tried very hard to use a consistent terminology throughout the document, despite text from seven different authors. In particular, the requirement level for every feature is described using one of the words: *must* (required), *should* (recommended), or *may* (optional). The word *must* is only used when support for a feature is required. *Should* has typically been used when the working group believed that hosts ought to conform to a rule, but felt there were (rare) valid situations in which the rule might be ignored. Finally, there are some features the committee considered worthy of mention but completely optional. These features are flagged with the word *may*. To give an example, if the document says:

“A host **must** support the KOM conference feature...”

then no implementation can claim conformance to the host requirements RFC without supporting this feature. But if the document says:

“Hosts **may** support the KOM conference feature...”

then a conformant host need not support the feature.

Link layer

The document opens with a short section on the link layer. Most link layer issues were discussed in RFC 1009, and so, for most issues, the host requirements specification simply references RFC 1009. While looking at this section, the working group concluded that RFC 1009 needs updating in some areas, and there are plans to revise RFC 1009 after the host requirements document is completed.

IP and ICMP

The section on the Internet layer protocols (IP and ICMP) follows the link layer. Much of this section is simply cleaning up nits: listing the five address classes (A through E), reminding implementors (yet again!) about proper use of broadcast addresses, the need to support subnetting, and the proper interpretation of the source route option.

The section also looks at key architectural issues, for example the host/gateway distinction. Hosts and gateways play very different roles in the Internet architecture (corresponding to End Systems and Intermediate Systems, respectively, in the ISO world.) The working group recognizes that there may be legitimate reasons for certain host systems to also act as gateways, i.e., to forward datagrams; this is called “embedded gateway functionality,” as in RFC 1009. However, the gateway function, if present, must follow RFC 1009, while the (“pure”) host function is described in the Host Requirements document. For example, the working group wants to discourage the practice of permitting (or requiring) all hosts to eavesdrop on gateway-to-gateway routing protocols.

continued on next page

The Host Requirements Working Group (*continued*)

Multihoming

The IP section also is the first one to touch a problem that crops up throughout the document: *multihoming*. Quite frankly the Internet architecture does not provide good support for multi-homing. As a result, the working group has had to examine the problem in several places. At the IP level, the working group had to worry about whether the IP source address of an outbound datagram had to match the IP address of the outbound interface (it does). At higher levels, the document looks at questions of whether applications have to be able to try second or even third addresses if they cannot connect to the first listed address (they do). Multihoming is a difficult problem, and the working group does not claim to have solved it.

UDP and TCP

Like the IP section, most of the UDP and TCP sections are devoted to clarifications of the specifications. For example, in the TCP section the document corrects various minor errors in the state diagrams and pseudo-code implementation. The sections also examine various common implementation practices, such as TCP Silly Window Syndrome avoidance (a good thing), delayed TCP acknowledgments (a mildly controversial feature), and turning off the UDP checksum (a very controversial feature).

Applications

Finally the requirements document looks at applications and support services: SMTP/RFC 822, Telnet, FTP, TFTP, the Domain Name System, network booting and network management. Various problems are cleared up. The Sorcerer's Apprentice bug in TFTP is explained and the recommended fix is given. Use of WKS records during MX lookups in SMTP is no longer required. The Telnet end-of-line issue is settled. The list of required FTP commands to be supported is expanded.

Checklist

At the end of the document there is an appendix containing a checklist of features. This checklist serves as a handy index to the document and should make it easier for vendors and buyers to determine if a system conforms to the specification. However, the ease of using the checklist contains a serious danger that it will be misused, and the working group included the checklist with some trepidation. Individual checklist entries are sometimes (deliberately) obscure, and can be interpreted only with the full text of the document. More importantly, the explanation and commentary of the text are essential to understand the significance of different checklist items. Some requirements are much more important than others, and some environments may justify deviations from the standard. Thus, it is our intent that the checklist be only a handy summary and index to the full text, and *never* be used by itself.

New work

To finish up the discussion of the host requirements effort, we should point out that the working group has also stimulated a few new RFCs. While preparing the requirements document, the group often found itself discussing new ways to fix well-known problems in the Internet architecture. Some of these discussions have generated their own RFCs. (The working group is very reluctant to propose untested ideas in the specification).

The current list of RFCs issued or expected based on the host requirements effort are:

- *RFC 1063: IP MTU Discovery Options*
This RFC explains a method, using two new IP options, for finding the MTU of a path through an IP internet. Use of these options will allow hosts to avoid fragmentation-related problems.
- *Gateway Discovery*
This document describes a method for hosts to learn who their default gateways are.
- *Telnet Terminal Type Negotiation*
An improved mechanism for using the option is explained.
- *RFC 1071: Computing the Internet Checksum*
A paper which describes the algorithms used to develop fast implementations of the Internet one's-complement checksum.

The Host Requirements Working Group has nearly completed its task, and the RFC is expected to be issued before the end of 1988.

CRAIG PARTRIDGE received his B.A. from Harvard University in 1983, and has been a part-time Ph.D. candidate there since 1987. For the past five years he has worked for BBN Systems and Technologies Corporation (formerly BBN Laboratories) on a variety of networking related projects including CSNET, the NSF Network Service Center (NNSC), and various projects concerned with distributed systems, IP transport protocols, and network management. In addition, he is a member of the Internet End-To-End Task Force, the Internet Engineering Task Force, and the Distributed Systems Architecture Board Task Force on Naming. He currently splits his time at BBN between CSNET, the NNSC, and managing a small TCP/IP networking project. Craig is also the editor of ACM SIGCOMM's *Computer Communication Review*.

BOB BRADEN has spent nearly a lifetime in the computer field, starting with 0th generation computers in 1951. In 1986 he came to ISI from UCLA, where he spent the preceding 18 years working for the computer center. He first became involved with the Arpanet in 1970, when his systems group made the UCLA 360/91 an Arpanet host (1/1, now 10.1.0.1!). He participated in design of the Arpanet FTP and RJE protocols and in TCP development, and wrote the UCLA MVS TCP/IP code. He is a charter member of the IAB and is chairman of the End-to-End Task Force. Bob definitely qualifies as an *Internet Old Boy*.



TCP in a bottle

From time to time we devote a little column space to the unusual or bizarre. Our feeling is that without a little humor and nonesense this publication would be very dry indeed. When our friends in the UK say "TCP" they are usually thinking about a multi-purpose liquid wonder-drug and not a transport protocol,—even if TCP/IP is popular in Europe. TCP is also available as lemon flavoured "Throat Pastilles," perfect for this time of the year. According to the package, TCP brand liquid is an aqueous solution of Phenol B.P. 0.175% w/v, halogenated phenols 0.68% w/v and Sodium Salicylate B.P. 0.052% w/v.

Improving Your TCP: Tuning the Checksum routine

by Craig Partridge, BBN Systems and Technologies Corp.

Introduction

Most TCP implementations spend the majority of their time computing *checksums*. This computation is expensive because it is the only TCP-level operation that requires examining every byte of data in the TCP segment. Because computing checksums is expensive, it is important to make sure that your routine is as efficient as possible. Indeed, tuning your checksum routine may be more important now than ever before. At the August SIGCOMM Conference, Van Jacobson explained how to implement TCP so that processing each inbound segment requires only fourteen (14) instructions beyond those in the checksum routine!

RFC 1071

A new RFC, RFC 1071 "Computing the Internet Checksum" by Bob Braden, Dave Borman and myself, was written to provide guidance to implementors trying to write fast checksum routines. (RFC 1071 also includes, as an appendix, IEN-45, an earlier work on the Internet checksum by Bill Plummer). Anyone actually tuning their checksum routine should consult this RFC. But, to help you figure out whether your checksum routine could use some tuning, the rest of this article presents an overview of the important techniques.

Hints

Essentially, the problem of computing the Internet checksum can be reduced to finding a way to quickly compute the *one's complement sum* of a sequence of 16-bit integers. The most important techniques for quickly computing the one's complement sum are:

- *Exploit byte-order independence of the sum.* It turns out that the sum is byte-order independent. If you compute the checksum in big-endian byte order and little-endian byte order, one sum will be the byte-swap of the other. As a result, you can always compute the checksum in your host's local byte order. Further, it doesn't matter in what order you add the integers in the sum.
- *Exploit bigger word sizes.* If your host has a word size that is larger than 16 bits, you can use a couple of techniques to speed up the sum.

One trick is accumulate the one's complement carries. One nuisance of the one's complement sum is that you have to add the carry back in. With a larger word size, you can accumulate the carries in the high-order bits of the word, and only add the carries back in at the end of several additions.

Another trick is to add in "parallel." It turns out to be possible to add 32 (or even more) bits at a time when computing the sum. After you've computed the 32-bit sum, you simply fold the two 16-bit halves of the word together to get the 16-bit sum.

- *Unroll the checksum loop.* It is usually advantageous to unroll the additions and do several additions in the inner checksum loop.

- *Combine with data copying.* Most systems copy data from one memory buffer to another at least once in the process of getting data from the application to the I/O board. Sometimes this data copy loop can be combined with the checksum routine for substantial performance gains.

If any of these techniques sound like they might benefit your checksum routine, it is probably worth your while to spend an hour reading RFC 1071.

Recent Publications

In our July issue, we had a list of information sources for TCP/IP. Since then several more have appeared:

SIGCOMM 88

The SIGCOMM 88 Symposium was held at Stanford University in August. The theme was Communication Architectures and Protocols. A number of Internet related papers were presented, in particular Van Jacobson's much awaited thesis on congestion avoidance and control. The proceedings are available from ACM Press at 301-528-4261.

New Books

Innovations in Internetworking edited by Craig Partridge and published by Artech House is a collection of 37 research papers written over the past twenty years. The goal of the collection was to pull together some of the most important research papers on problems in each of the seven layers of the OSI (inter)network model. Such a collection would be suitable both as a reference collection on a computer scientist's bookshelf, and also as a book of readings for a graduate seminar. The papers in the collection come from research with a variety of internetwork architectures including DECNet, TCP/IP, OSI and XNS. The ISBN number is 0-89006-337-0. For more information call Artech House Books at 800-225-9977 x4002.

Internetworking Computer Systems, John McConnell, Prentice-Hall, ISBN 0-13-473976-0

Internetworking—A Tutorial on Bridges and Routers, Available from 3Com Corporation, Santa Clara, CA, 408-562-6400.

Internetworking—An Introduction, Available from The Wollongong Group, Palo Alto, CA, 415-962-7100.

Bibliography

The Experimental Literature of The Internet: An Annotated Bibliography, Jeffrey C. Mogul, Digital Equipment Corporation, WRL Technical Note TN-1, August 1988. Abstract: "The DARPA Internet is the most successful experiment in heterogeneous internetworking. It connects computer systems from almost every major vendor, using a wide variety of wide-area and local-area network technology, and is in continual use by thousands of people. This annotated bibliography covers the literature of the Internet as an experiment: publications which convey the experience acquired by the experimenters." TN-1 is available from:

Technical Report Distribution
 DEC Western Research Laboratory, UCO-4
 100 Hamilton Avenue
 Palo Alto, CA 94301
 E-mail: WRL-Techreports@decwrl.dec.com

CONNEXIONS
480 San Antonio Road
Suite 100
Mountain View, CA 94040

FIRST CLASS MAIL
U.S. POSTAGE
PAID
SAN JOSE, CA
PERMIT NO. 1

CONNEXIONS

PUBLISHER Daniel C. Lynch

EDITOR Ole J. Jacobsen

EDITORIAL ADVISORY BOARD Dr. Vinton G. Cerf, Vice President, National Research Initiatives.

Dr. David D. Clark, The Internet Architect, Massachusetts Institute of Technology.

Dr. David L. Mills, NSFnet Technical Advisor; Professor, University of Delaware.

Dr. Jonathan B. Postel, Assistant Internet Architect, Internet Activities Board; Division Director, University of Southern California Information Sciences Institute.

Subscribe to CONNEXIONS

U.S./Canada \$100. for 12 issues/year
International \$ 50. additional per year

Name _____ Title _____

Company _____

Address _____

City _____ State _____ Zip _____

Country _____ Telephone () _____

Check enclosed (in U.S. dollars made payable to CONNEXIONS). Bill me/PO# _____
 Charge my Visa Master Card Card # _____ Exp. Date _____

Signature _____

Please return this application with payment to:
Back issues available upon request \$10./each

CONNEXIONS
480 San Antonio Road Suite 100
Mountain View, CA 94040
415-941-3399